

Two decades after Chao's work, Li and Thompson (1977) conduct the first systematic study in acquisition of tone of Mandarin. They delineate four stages in the tone acquisition process, shown as followings. (1977:189)

Stage I: The child's vocabulary is small. High and falling tones predominate irrespective of the tone of the adult form.

Stage II: The child is still at the one-word stage, but he has a larger vocabulary. The correct 4-way adult tone contrast has appeared, but sometimes there is confusion between rising and dipping tone words.

Stage III: The child is at the 2/3-word stage. Some rising and dipping tone errors remain. TS is beginning to be acquired.

Stage IV: Longer sentences are being produced. Rising and dipping tone errors are practically non-existent.

In addition, they also find that (1) the correct tone system is acquired relatively quickly and is mastered well in advance of the segmental system; (2) the high and falling tones are acquired earlier and more easily than the rising and dipping tones; (3) substitution between the rising and dipping tones persists throughout stages II and III; (4) the TS rules are learned, with infrequent errors, as soon as the child begins to produce his own multi-word utterances (1977:189).

Another important work is Tse's (1978) longitudinal case study on the tone acquisition in Cantonese, in which he hypothesizes several universal principles for the chronological order of tone acquisition: (1978:32)

- (1) The acquisition of level tones precedes that of contour tones.
- (2) With contour tones, the acquisition of falling tones precedes that of rising tones.
- (3) The acquisition of the high rising tone precedes that of the low rising tone, in a language where there are two rising tones.

Besides the acquisition order, Hsu (1987) provides answers to duration for tone acquisition from his cross-sectional longitudinal study on twenty-eight Mandarin-speaking children for three years:

- (1) The first tones to appear are T1 and T4 at the age of 1;0, but there are some mutual substitutions.
- (2) All T2 and T3 are substituted by T1 or T4 before 1;0.
- (3) T3 emerges sporadically at around the age of 1;6.
- (4) By the time of 2;6, all the subjects have acquired the four tones, but there is still confusion between T2 and T3.
- (5) All the subjects have acquired TS1 and TS2 before 3;0. (1987:89)

Therefore, the duration of the process of tone acquisition in Hsu's study is eighteen months; and another six months is required for the tone sandhi rules to be acquired.

Xu (1989) in her thesis confirms Tse's hypothesis of universal principles, namely the precedence of level tones and falling tone. As for TS, Hsu's finding also supports Li and Thompson's generalization that the TS rules are learned, with infrequent errors, as soon as the child begins to produce multi-word utterances. As for tone acquisition, Xu ranks the accuracy of four tones in the ordering as follows: tone1>tone4>tone2>tone3. It is also indicated that in the span of time for two years, subjects had not completed the acquisition of the Mandarin tonal system, which shows disagreement with previous studies where a duration of seven months in Li and Thompson (1977:279) and the time span of eight months by Tse (1978:198).

2.1 Literature review on the role of input

After reviewing studies on tone acquisition, valuable studies on the effect of adult's input in L1 acquisition will be presented. While most of the previous studies investigated English-speaking children's phonological acquisition, very few papers presented the role of adult's input in children acquiring tonal system in Mandarin.

Skinner in *Verbal Behavior* (1957) views language learning as the result of a habit-forming process from positive or negative reinforcement. Supported by the success of experimental studies in which features hypothesized to be helpful were actually shown to promote syntactic acquisition provided incontrovertible evidence in support of the value of appropriate input (Baker and Nelson 1984), behaviorism was enthusiastically received. Specifically, the most powerful input indicated in the

training experiments were recasts: responses to child utterances that provided corrected or alternative versions, showing clearly feedback are utilized by children. In addition, several studies have discovered that parents do respond differentially to the grammatical and ungrammatical utterances produced by their children. These differential responses, including corrections, could be enough to signal an error to the child by repeating more of their children's ill-formed utterances than well-formed one. (Hirsh-Pasek, Treiman, and Schneiderman 1984)

It has been shown that adult speech is often assumed to be immediate input for a child's linguistic analysis. However, adult input has to be comprehended by the child before its linguistic properties can be analyzed. There are authors who rightly point out that a major problem with comprehension as a source of input for children's analysis is that the speech signal fades so quickly; therefore, it does not seem reasonable to assume that the language learner can somehow apprehend the fast-fading message produced by someone else, figure out what it means and how it is put together, and then relate it to similar utterances he has heard.

In view of this problem, Elbers (1989) argues that if comprehension does not qualify as a likely source of input for analysis, production does. His output-as-input hypothesis is based on a three-phase, inductive acquisition model, which takes the child's own productions as primary input for analysis. In Phase A, the child takes in incompletely analyzed fragments from adult input, and starts using these in her or his own production. In Phase B, own productions are analyzed and hypotheses concerning relationships between forms and meanings are derived. In Phase C, these hypotheses are tested against novel adult input.

It will be proved by the result of this paper that adult's input does influence greatly on children's language acquisition. Specifically, the output-as-input hypothesis explains best children's TS rules acquisition in Mandarin.

2.2 Expectations

Based on the literature reviews, expectation of children's acquisition of tone and tone sandhi is described as follow. (Li & Thompson's four stages are adopted)

Stage I (1,0): before segments acquisition, T1 and T4 are acquired. Specifically, T1 appears earlier than T4 for the reason that level tone is acquired earlier than contour tone.

Stage II/one-word stage (1,6): T2 and T3 appear in the sequence of T2>T3. However, substitution between these two tones occurs.

Stage III (2/3 word stage) (2,4): Confusion between T2 and T3 remains. All four tone sandhi appear with the order: TS3>TS2/TS4>TS1.

Stage IV (sentence stage) (3,0): Both TS2 and TS1 appear stably and confusion between T2 and T3 no longer exists.

Though no previous studies mention the acquisition of TS3 and TS4, it will be discussed in the paper. The expectation for TS3 is that TS3 should appear at stage III earlier than the appearance of TS2 because TS3 only involves early-acquired T4. And the expectation for TS4 is that TS4 should appear when children start combining these two lexemes with other lexemes, on early period of stage III. Therefore, the expectation of tone acquisition is T1>T4>T2>T3 as Xu suggested. And the expectation of TS acquisition is TS3 > TS2/TS4 > TS1, and all at stage III.

As for the role of adult's input, a close examination will be made during the observation. If Baker & Nelson's immediate recasts are correct, it is expected that subjects would more or less modify their production after given immediate recasts. If Elber's output-as-input hypothesis is correct, it is expected that subjects would improve only those they choose to produce after given immediate recasts. And if Innatism theory is correct, it is expected that adult's input does not help acquisition.

3. Methodology

Three children aged respectively (1;5), (2;2) and (3;0) were observed. They have Mandarin as their first language, and have very little influence of any second language. The period of data collection is two months, from April to May, 2006.

Data collection on tone acquisition was done by 6-hour-videotaping when subjects were playing with their parents. Error-counting on recorded tokens was done to decide at which acquisition level the

subject is. As for the procedure of TS acquisition, we engaged the subjects in looking at pictures. The adult pointed to the picture and asked the child what it was, and tried to elicit a naming response from the child. All the collected tokens in picture-naming session were analyzed targeting to answer research question (1) ‘is there an ordering in the acquisition of four TS rules under the influence of tone acquisition?’

After each utterance, the adult would provide a recast as correct model and asked the subject to repeat. This repetition is to evaluate the influence of adult’s input targeting to answer the second research question, ‘how does adult input help acquisition of TS rules?’

4. Result/discussion

Table 1 presents the correct usage of tones and TS rules of three subjects.

Table 1 Number and Accuracy rate of three subjects

	Age (1,5) Correct/Production tokens	Age (2,2) Correct/Production tokens	Age (3,0) Correct/Production tokens
T1	30/37 (81.08%)	185/191 (96.86%)	356/356 (100%)
T2	43/74 (58.11%)	225/249 (90.36%)	388/388 (100%)
T3	30/73 (41.09%)	181/363 (49.86%)	167/167 (100%)
T4	27/44 (61.36%)	432/442 (97.74%)	464/464 (100%)
TS1	0	2/6 (33.33%)	6/6 (100%)
TS2	0	4/6 (66.67%)	6/6 (100%)
TS3	0	6/6 (100%)	6/6 (100%)
TS4	0	5/6 (83.33%)	6/6 (100%)

The (1,5) subject, who is at Stage II, is expected to produce stably T1 and T4, and T2 and T3 appear in the sequence of T2>T3. Substitution between these two tones occurs. However, data collection from this subject is difficult. All tokens in the videotaping and testing sessions are elicited, for this child rarely spontaneously utters any speech. Therefore, table 1 shows the accuracy rate of the (1,5) subject’s *repeated* production.

Accuracy percentage of tone production of (1,5) subject is shown as T1>T4>T2>T3. T3 is often pronounced wrongly. And in these wrong pronunciation, the subject always substitutes T2 for T3, such as *hand* ‘手’ as [sou2] and *yes* ‘好’ as [hau2]. It is only shown once in our observation that the subject substitutes T3 for T2 in *魚* as [yu3]. This asymmetry tells that T3 is more difficult than T2, and disagrees with previous papers saying that T2 and T3 substitute each other, for a clearly greater difficulty of T3 for the (1;5) subject.

It is interesting to find out that the accuracy rate of T4 is only slightly higher than the accuracy rate of T2. Based on the expectation, this (1;5) subject is at stage II where T1 and T4 should appear quite stably while T2 and T3 much less stably. However, the result suggests that T4 may be acquired roughly at the time when T2 is acquired, for both tones are slanting tone. However, further studies are acquired to explore into this preliminary observation.

In addition, the observation of this (1;5) subject also clearly tells us that the tonal system is acquired earlier than the acquisition of phonemes. When the subject repeats, most the tone is produced even though the phonetic configuration is not correct. For example, *鞋* ‘shoes’ as [ie2], and *樹* ‘trees’ as [u4]. It is shown that the tonal system is acquired earlier than the phoneme acquisition.

The observation of (2;2) child clearly provides an answer to the question of the order of tone sandhi acquisition under the influence of tone acquisition. As can be seen, in the 6-hour-videotaping, all four tones appear. T1 and T4 are stable, while T3 is relatively unstable. Based on the percentage of accuracy, we may order a difficulty scale which supports the previous studies on tone acquisition, stating that T1 and T4 are easier to acquire than T2 and T3; particularly, T3 is the most difficult one, shown by a low percentage of accuracy.

The expectation for (3,0) subject, who is at stage IV, is that confusion between T2 and T3 no longer exists. Observation of 3-year-old subject accords with the expectation. The subject produces four tones

without any confusion both in testing and videotaping sessions. We then conclude that the (3;0) subject has completed acquiring tones in Mandarin.

Let's move to the TS acquisition. For the (1,5) subject, who is at one word stage, his mother tries to elicit the repetition of some frequently used terms, such as 謝謝 'thank-you', 再見 'goodbye', 好吃 'good taste', and 手錶 'watch'. It is found that only TS3 is applied in 謝謝 'thank you' and 再見 'goodbye'. Though the data is limited, the appearance of TS3 accords with the expectation that TS3 is the first TS rule to be acquired.

At the age of (2;2), it is observed that all four TS rules appear with different proficiency. TS3 (T4 → high falling tone/ _ T4) appears most stably. The next stably-applied rule is TS4, then TS2, and the least stably TS1. The accuracy percentage appearing in testing sessions is supported by the accuracy percentage of subject's applying tone sandhi rules in the videotaping session. Table 2 shows the percentage of accuracy in the videotaping session of (2,2) subject.

Table 2 Percentage of accuracy in applying TS rules in videotaping session of (2,2) subject

	TS1	TS2	TS3	TS4
Total appearance	47	86	101	35
Correct appliance	16	65	96	35
Percentage of accuracy	34.04%	75.58%	95.05%	100%

Table 2 supports Table 1 in the fact that it is clearly seen that TS3, which has high accuracy rate in testing session, also has high accuracy rate in the videotaping session. For example, in the production of bisyllabic or trisyllabic T4 phrases, such as 樹上 'in the tree', 去睡覺 'go to bed', the subject applies TS3 without mistakes.

As for TS1, the accuracy percentages both in the testing session and the videotaping session are the lowest. Errors, such as 給我 'give me' as [ge1 wo3] and 給你 'give you' as [ge1 ni1] are observed despite the subject can correctly pronounce the isolated morpheme. This shows that the subject is already aware that a special attention should be given to the adjacent dipping tones; however, for the reason that appliance of the Third Tone Sandhi Rule has not been fully acquired, the subject fails to produce the correct tones. Even so, the subject shows her own ways to avoid the adjacency of two third tones.

The expectation for (3,0) subject at Stage IV is that all TS rules are applied correctly. Indeed, the subject applies all four TS rules correctly when needed. The stably acquired tonal system is supported by the videotaping session, where the accuracy rate for four TS rules are nearly 100%. We then conclude that the (3;0) subject has completed acquiring tonal system in Mandarin.

In conclusion, from the observation of three subjects, the results generally accord with our expectations. Specifically, the tone acquisition order and TS rules acquisition order are the expected T1>T4>T2>T3 and TS3/TS4>TS2>TS1, which shows that the order of TS acquisition is influenced by tone acquisition. It is also shown that not until the age of three does the subject complete acquiring tonal system in Mandarin. Though within the short period of two months of observation, we cannot tell the specific duration the subject acquires the tone or tone sandhi, the difficulty sequence in acquisition observed from (1,5) and (2,2) subjects provides the clue. T1 and T4 are easier than T2, which is easier than T3. As for TS acquisition, TS3 and TS4 are easier than TS2, which is easier than TS1.

After answering research questions (a), in the next session, adults' input will be discussed.

4.4 Influence of adult's Input

Since the three-year-old subject has acquired the tonal system completely, no repetition task is necessary, we shall only discuss the influence of adult's inputs on (1,5) and (2,2) subjects.

For the reason that all data from (1,5) child is from elicitation, it is clear that parents' input influences the production. For example, the first repeated production of 好吃 'good taste' is [hau1 tɕʰɿ1], and the second repeated production is the correct tones [hau3 tɕʰɿ1]. However, the imitation is not an easy task. In the production of 手錶 'watch', the first repetition is [sou1][biau2], and the second repetition is [sou3][biau2], and the third repetition is [sou1][biau3]. The subject seemed to be able to focus on only one sound at one time.

As for the (2,2) subject, parents input also greatly influences her tonal production. When she said 給我 ‘give me’ as [gai1 wo3] for the first time, but [gai2 wo3] after her mother’s input, and 給你 ‘give you’ as [gai1 ni1] for the first time, but [gai2 ni3] after adults’ input. Not only the sandhi tones are modified, but also tones are modified. For example, 椅 ‘chair’ is pronounced with low level tone for the first time, but correctly pronounced after repetition.

Though adult’s input seems to have immediate effect, we want to know if, or specifically how, adult’s input really help the acquisition. Therefore a follow-up observation of (2;2) subject three weeks after the testing session is made to see if the subjects can apply correct TS rules after given recasts in testing session.

Interestingly, there is a close correlation of successful acquisition examined three weeks after the testing session and simultaneous repetition right after the recasts in testing session. It is observed that immediate input has effects only when the child’s spontaneous-repetition after the input occurs. To analyze the relationship between acquisition and simultaneous repetition, subjects’ input utterances are classified as spontaneous or non-spontaneous and in terms of whether they were responded to or not. Table 3 shows the result.

Table 3 Children’s responses to input

	TS1	TS2	TS3	TS4
Spontaneous reaction	6	5	0	0
Non-spontaneous reaction	0	1	2	0
No reaction	0	0	4	6

To see if immediate input has effects, we have to see if the subject performs better in appliance of TS rules after three weeks. Table 4 shows the number and percentage of accuracy of the follow-up test.

Table 4 Number and accuracy rate of the follow-up test

	TS1	TS2	TS3	TS4
Number of test items	6	6	6	6
Correct appliance	5	6	6	5
Accuracy percentage	83.33%	100%	100%	83.33%

A comparison of table 4 and table 1 is made to see the subject’s improvement and effects of adult’s input, shown in figure 1.

Figure 1 Comparison of accuracy rate of first and follow-up tests

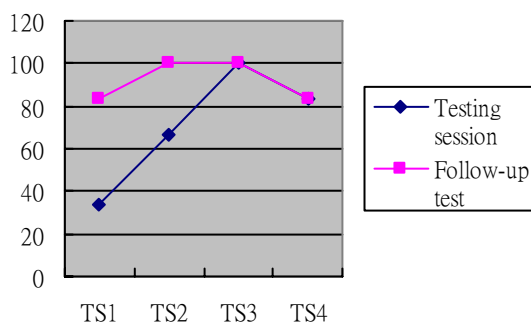


Figure 1 clearly shows the improvement of correct appliance of TS1 and TS2, and this improvement correlates with the simultaneous repetition in testing session shown in table 3. This suggests that when the mother takes the opportunity to provide the child with the correct realization of tones, children do not always react to the immediate input. Only when children spontaneously repeat adult’s input, it is more likely the child learns the correct realization.

This finding has several implications. Firstly, there is no a priori reason to expect that, at any given time, a child is ready to acquire the element being modeled. If the child is not ready, then no amount of feedback will lead to immediate acquisition.

Secondly, the result shows support for the output-as-input hypothesis, for only when the subject immediately repeats the recasts does the input effects child's acquisition process. It seems that the subject selects the input which is comprehensible in his/her acquisition level. However, although we know that children severely deprived of input and that children do not learn language from mere exposure alone, there are indications that because of the recasts, subjects have the opportunities to practice the sound pattern.

Next question to be discovered is that why the subjects choose certain input to practice. A question is expected to study further by other researchers. It is acknowledged that children are selective listeners who may or may not be prepared to attend to particular adult speech; investigators have very little understanding of the reason. However, within Elbers' output-as-input model, a child's readiness may be explained. It is generally observed that children's attention tends to be attracted most by that which is slightly different from expectation, that is, by that which is slightly different from what they would predict on the basis of acquisition level.

This output-as-input model points out three crucial features observed from the TS acquisition of this paper

1. Adult input is not unimportant. Rather, the model specifies what exactly adult input is important for. Adult input is held to be essential for intake and testing, but not for analysis, which is primarily dependent on own production.
2. In the output-as-input model, it is not only the case that production draws on perception, but also that perception and selective attention are influenced and guided by prior production.
3. A final strong point of the model is the ease with which it can accommodate to the acquisition of TS rules in Mandarin.

5. Conclusion

The observation of three subjects (1;5, 2;2, 3;0) accords with the expectations of acquisition of tone and TS in Mandarin. In tone acquisition, the order of difficulty is $T1 < T4 < T2 < T3$. Different from previous researches who suggest that at stage II where T1 and T4 are acquired stably and substitution of T2 for T3 occurs, it is found that T4 is acquired at a time when T2 is acquired and the subject only uses T2 to substitute T3 but not vice versa. At stage III, all four TS rules appear with the order: $TS3/TS4 > TS2 > TS1$. And by the age of three or stage IV, tone and TS rules are fully acquired.

Back to the two research questions related to the TS acquisition:

- (a) Is there an ordering in the acquisition of four TS rules under the influence of tone acquisition?
- (b) How does adult's input help acquisition of tone sandhi rules?

The answer to question (a) is 'yes, the order of tone acquisition decides the order of TS acquisition'. It is also observed in (1;5) subject that tonal system is acquired earlier than the segments which accord with previous studies, and in (3;0) that not until the age of three have children acquired the tonal system of Mandarin completely, which disagrees with Li and Thompson's quick acquisition of seven months, and Tse's observation of eight months.

Now, we may turn to the other research question of this paper which is to examine the role of adults' input in tone sandhi acquisition. Based on the result of the follow-up test session, we argue that the output-as-input hypothesis correctly explains the correlation between successful acquisition and subject's simultaneous repetition to adult's input. We also argue for the advantage of adopting this hypothesis as the model understanding the children's selective response to certain input and also a great contribution of this hypothesis: that is this hypothesis acknowledges the importance of adult's input in acquiring tone and TS rules.

6. Further studies

For the limited period of time for observation, and limited number of subjects, this paper cannot provide a definite answer to the question about the duration of TS rules acquisition. It is hoped that future studies may conduct a well quantitative or qualitative study to help us have a better understanding about tonal system acquisition in Mandarin.

References

- Chao, Y. R. 1951. The Cantian idiolect: an analysis of the Chinese spoken by a twenty-eight-month-old child. *Studies of child language development*, ed. by Charles A. Ferguson & D. I. Slobin. New York: Holt, Rinehart and Winston, INC.
- Chao, Y. R. 1968. *A grammar of spoken Chinese*. Berkeley: University of California Press.18-56.
- Elbers, Loekie. 2000. An output-as-input hypothesis in language acquisition. *Models of Language Acquisition: inductive & deductive approaches*, ed. by Peter Broeder and Jaap Murre. Massachusetts: Oxford University Press.244-71.
- Gleason, Jean Berko. 1978. Talking to children: some notes on feedback. *Talking to children: language input & acquisition*, ed. by Catherine E. Snow and Charles A. Ferguson. Cambridge: Cambridge University Press.199-205.
- Hirsh-Pasek, K., Treiman, R., and Schneiderman, M. 1984. Brown and Hanlon revisited: mother's sensitivity to ungrammatical forms. *Journal of Child Language* 11.81-8.
- Ioup, Georgette. Testing the relationship of formal instruction to the input hypothesis. *TESOL quarterly*.345-50.
- Li, Charles & Thompson, Sandra A. 1977. The acquisition of tone in Mandarin-speaking children. *Journal of Child Language* 4.185-99.
- Sokolov, Jeffrey L. & Snow, Catherine E. 1994. The changing role of negative evidence in theories of language development. *Input and Interaction in language acquisition*, ed. by Clare Gallaway and Brian J. Richards. Cambridge: Cambridge University Press.38-55.
- Tse, K.-P. John. 1978. Tone Acquisition in Cantonese: a Longitudinal Case Study. *Journal of Child Language* Vol.5.191-204.
- Xu, Hui-xiu. 1989. 兒童國語音韻習得:九個月至三歲個案研究. MA thesis in National Taiwan Normal University.
- Zhu hua & Barbara Dodd. 2000. The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language* 27. Cambridge: Cambridge University Press. 3-42.